



Systems biology analysis of hepatitis C virus infection reveals the role of copy number increases in regions of chromosome 1q in hepatocellular carcinoma metabolism

Elsemman, Ibrahim; Mardinoglu, Adil; Shoaie, Saeed; Soliman, Taysir H.; Nielsen, Jens

Published in:
Molecular BioSystems

Link to article, DOI:
[10.1039/c5mb00827a](https://doi.org/10.1039/c5mb00827a)

Publication date:
2016

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Elsemman, I., Mardinoglu, A., Shoaie, S., Soliman, T. H., & Nielsen, J. (2016). Systems biology analysis of hepatitis C virus infection reveals the role of copy number increases in regions of chromosome 1q in hepatocellular carcinoma metabolism. *Molecular BioSystems*, 12(5), 1496-1506.
<https://doi.org/10.1039/c5mb00827a>

General rights

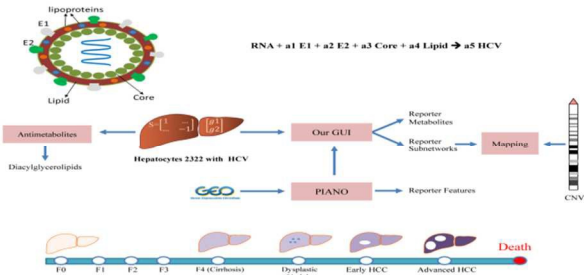
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Table of contents entry

We reconstructed Hepatitis C Virus assembly reactions to find host-target metabolites impeding this reaction.



**Systems biology analysis of hepatitis C virus infection reveals
the role of copy number increases in regions of chromosome 1q
in hepatocellular carcinoma metabolism**

Ibrahim E. Elsemman^{1, 2,5}, Adil Mardinoglu^{1,3}, Saeed Shoaie¹, Taysir H. Soliman⁴, and Jens Nielsen^{1,3*}

¹Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

²Department of Mathematics, Faculty of Science, Assiut University, Assiut, Egypt

³Science for Life Laboratory, Royal Institute of Technology, Stockholm, Sweden

⁴Information Systems Department, Faculty of Computers and Information, Assiut University, Assiut, Egypt

⁵ The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark

* Corresponding author

E-mail: nielsenj@chalmers.se

1 **Abstract**

2 Hepatitis C virus (HCV) infection is a worldwide healthcare problem; however, traditional
3 treatment methods have failed to cure all patients, and HCV has developed resistance to new
4 drugs. Systems biology-based analyses could play an important role holistic analysis of the
5 impact of HCV on hepatocellular metabolism. Here, we integrated HCV assembly reactions with
6 a genome-scale hepatocyte metabolic model to identify the metabolic targets for HCV assembly
7 and the metabolic alterations that occur between different HCV progression states (cirrhosis,
8 dysplastic nodule, and early and advanced hepatocellular carcinoma (HCC)) and healthy liver
9 tissue. We found that diacylglycerolipids were essential for HCV assembly. In addition, the
10 metabolism of keratan sulfate and chondroitin sulfate was significantly changed in the cirrhosis
11 stage, whereas the metabolism of acyl-carnitine was significantly changed in the dysplastic
12 nodule and early HCC stages. Our results explained the role of the upregulated expression of
13 *BCAT1*, *PLOD3* and six other methyltransferase genes involved in carnitine biosynthesis and S-
14 adenosylmethionine metabolism in the early and advanced HCC stages. Moreover, *GNPAT* and
15 *BCAP31* expression was upregulated in the early and advanced HCC stages and could lead to
16 increased acyl-CoA consumption. By integrating our results with copy number variation analyses,
17 we observed that *GNPAT*, *PPOX* and five of the methyltransferase genes (*ASH1L*, *METTL13*,
18 *SMYD2*, *TARBPI* and *SMYD3*), which are all located on chromosome 1q, had increased copy
19 numbers in the cancer samples relative to the normal samples. Finally, we confirmed our
20 predictions with the results of metabolomics studies and proposed that inhibiting the identified
21 targets has the potential to provide an effective treatment strategy for HCV-associated liver
22 disorders.

Introduction

More than 185 million people (3% of the world population) are infected with the hepatitis C virus (HCV), which kills approximately 350,000 people every year. HCV infection leads to chronic hepatitis C, cirrhosis and hepatocellular carcinoma (HCC).¹ HCV complications have also been associated with type 2 diabetes² and coronary heart disease.³ Although HCV was discovered in the 1980s⁴, a vaccine is not available because of challenges associated with vaccine development, including the lack of a cell culture system for studying T cell responses.⁵ The traditional treatment strategy for HCV is effective in only 50% of patients, and its success depends on genetic variations of interleukin-28.⁶ Despite success with new interferon-independent HCV therapies¹, new treatment strategies for HCV are urgently needed because HCV has developed resistance to these new anti-HCV drugs.⁷ To overcome these challenges, systems biology-based approaches can play important roles in understanding the progression of HCV-associated liver disorders⁸, discovering novel biomarkers to detect HCV infection at an early stage and identifying novel drug targets for the development of effective treatment strategies.^{9, 10}

HCV is the primary cause of HCC in Europe and North America, whereas hepatitis B virus (HBV) is the primary cause of HCC in Asia and Africa.¹¹ Complications in the diagnosis of HCC at the early stage restrict its treatment.¹² One of the clinical biomarkers is α -fetoprotein (AFP), which has a low sensitivity for the detection of HCC caused by viral infections.¹³ Liver biopsies also play an important role in the early diagnosis of HCC but may lead to major complications, including pain, bleeding and death.¹⁴ Therefore, identifying non-invasive biomarkers is important¹², especially for the diagnosis of the dysplastic nodule stage of HCV.^{15, 16}

An alternative to invasive diagnostic tools is the application of HCC biomarkers that may be identified through the use of omics techniques where control samples are compared with

1 pathological samples.¹⁷ Transcriptomics data have been generated to identify significantly
2 differentially expressed genes between early HCC and the dysplastic nodule stage.¹⁶
3 Metabolomics studies have also been performed to identify metabolic biomarkers and reveal the
4 molecular mechanisms underlying HCC.¹⁸ Based on this candidate metabolomics biomarkers that
5 vary across serum, urine and tissue were identified for detection of HCC at the early stage¹⁹;
6 however, to our knowledge, they have not been used in clinical applications.

7 To understand how the molecular mechanisms within liver tissue change in response to HCV,
8 genome-scale metabolic models (GEMs) can be used as scaffolds for the integration of omics
9 data for a specific cell/tissue type and the elucidation of genotype–phenotype relationships.^{20–22}
10 GEMs include the known biochemical reactions that occur in a particular cell or tissue type, and
11 each reaction is associated with one or more gene(s).²³ GEM content is annotated from the
12 literature and bioinformatics databases, such as the KEGG database²⁴. These integrative models
13 enable studies of gene functions using a holistic approach. GEMs can also be used to interpret
14 high throughput clinical data using the network structure provided by these models.²⁵ For
15 instance, the Reporter Metabolite and Reporter Subnetwork algorithms can integrate
16 transcriptomics data with GEMs to explore metabolic “hot-spots” between healthy and diseased
17 states (i.e., metabolic functions that are altered because of different metabolic demands or to
18 maintain homeostasis).²⁶ Additionally, GEMs can be used to identify drug targets for human
19 pathogens^{27, 28} and cancer²⁹ using the antimetabolite concept, which blocks all fluxes around a
20 particular metabolite.³⁰ Recently, a comprehensive GEM for hepatocytes (*iHepatocytes2322*) was
21 reconstructed based on HMR 2.0³¹. This GEM includes the contents of the previously published
22 GEMs for liver tissue and an extensive description of lipid metabolism.³¹
23

1 Here, we used *iHepatocytes2322* to determine the metabolic interactions between HCV and
2 healthy hepatocytes (**Figure 1A**). First, we included the reactions associated with the HCV
3 assembly into *iHepatocytes2322* and employed the antimetabolite concept to identify the set of
4 metabolites that inhibit HCV development. Second, we used the Reporter Metabolite and
5 Reporter Subnetwork algorithms to integrate transcriptomics data with *iHepatocytes2322* to
6 reveal the molecular mechanisms underlying the progressive HCC stages (cirrhosis, dysplastic
7 nodule, and early and advanced HCC) caused by HCV infection. Through the Reporter
8 Metabolite analysis, we found that the metabolic functions (but not the concentration) of the
9 metabolites keratan sulfate and chondroitin sulfate were significantly changed in the cirrhosis
10 stage. Moreover, acyl-carnitine metabolism was significantly changed in the dysplastic nodule
11 and early HCC stages but not in the cirrhosis stage. These identified metabolites may be
12 considered candidate biomarkers for early HCC and can be measured in the blood or urine. Next,
13 we applied the Reporter Subnetwork algorithm and found connected reactions involved in the
14 progression of HCC. Finally, the integration of these target genes with the copy number
15 variations (CNVs) from The Cancer Genome Atlas (TCGA) revealed that a number of the genes
16 (i.e., *ASH1L*, *METTL13*, *SMYD2*, *TARBP1* and *SMYD3*) located on chromosome 1q were
17 amplified. Thus, our study revealed altered metabolism in response to HCC and identified
18 candidate biomarkers and therapeutic targets.

19

1 **Results**

2 **Reconstruction of the *iHCV***

3 A GEM can describe how a cell utilizes carbon, phosphate, and nitrogen sources as well as other
4 autotrophic metabolites, such as amino acids, to form its biomass.³² However, a virus do not have
5 internal metabolism required to build its structure, which contains a DNA or RNA genome and
6 structural proteins, although its structure may also contain a lipid raft. Thus, the viral structure is
7 assembled and replicated using the resources of the host. *In silico* viral models assume that the
8 virus is a pseudo-metabolite assembled from resources provided by the host model.^{33, 34} These
9 models contain a number of reactions that describe how viral genomes and proteins are
10 biosynthesized. For HCV (**Figure S1A**), a host ribosome requires tRNA to translate the HCV
11 genome, a single positive-strand RNA sequence, to three structural proteins (the core protein and
12 envelope glycoproteins *E1* and *E2*), and seven non-structural proteins (*p7*, *NS2*, *NS3*, *NS4A*,
13 *NS4B*, *NS5A* and *NS5B*).³⁵ The HCV RNA polymerase (*NS5B*) requires host RNA nucleotides to
14 transcribe the HCV RNA genome.³⁶ Furthermore, HCV requires a host lipid raft³⁷. The lipid
15 composition of HCV particles extracted from the plasma of 36 patients was measured together
16 with the triglycerolipid (TAG) per *apoB* and the *apoB* per HCV RNA³⁷ (**Table S1**). Here, we
17 assumed that HCV was a generic large metabolite that could be assembled from RNA, structural
18 proteins and lipids (**Figure 1A**).

19

20 To construct the HCV assembled reaction (*iHCV*), we downloaded the JFH1 sequence (a
21 Japanese HCV genotype 2a strain) from the UniProt web site using the entry ID Q99IB8. We
22 compiled the nucleotide composition in the HCV genome and the amino acid composition for
23 each HCV protein. From this information and the lipid composition measurements (**Table S1**),

we reconstructed the *iHCV* containing 13 reactions: ten reactions for the biosynthesis of HCV proteins, one reaction for HCV RNA genome biosynthesis, one reaction for the HCV lipid composition and one reaction for the assembly of HCV from the RNA genome as well as for the apoB and lipid composition (**Figure 1A**). Finally, *iHCV* was incorporated into *iHepatocytes2322* to generate the HCV liver model. We used the hepatic uptake rates from a mouse study for the model simulations.^{38, 39} Additionally, we used the uptake rates for remnants of chylomicrons, very-low-density lipoproteins (VLDLs), low-density lipoproteins (LDLs) and high-density lipoproteins (HDLs) as 0.0001 mmol/g liver/h.⁴⁰ The updated model is provided in SBML and Excel format (**Supplementary Data 1**).

Diacylglycerolipids (DAGs) are essential for HCV assembly

We used the antimetabolite concept to identify reactions that are essential for HCV assembly. Antimetabolites are the structural analogs of metabolites, and they are used to block the reactions that consume their corresponding metabolites (**Figure S1B**). The functions of antimetabolites can be simulated using GEMs by maximizing the flux balance analysis (FBA) objective function (here, HCV assembly) when all of the fluxes using a given metabolite are set to zero.²⁷ Agren et al⁴¹, Ghaffari et al²⁹ and Kim et al^{27, 28} used this method to identify targets for HCC, cancer cell lines and human pathogens^{27, 28}, respectively. Here, we applied this method twice to indicate when the model can and cannot consume VLDL, HDL and LDL remnants from the extracellular environment because the HCV lipid compositions are similar to VLDL and LDL⁴² and the hepatocytes can consume all of these remnants³¹. The antimetabolites included amino acids, nucleotides and cofactors that are essential for hepatocytes; therefore, we discarded these antimetabolites. Additionally, when certain metabolites are removed from the model, it is not possible to find a solution using FBA. We also discarded these metabolites because there were no

flux distributions to synthesize the essential metabolites for hepatocytes and HCV, such as amino acids, RNA nucleotides and lipid metabolites.

Without the uptake of the VLDL, LDL and HDL remnants, we identified 11 sterol metabolites, such as squalene, 3-dehydrosphinganine, sphinganine, 1,2-diacylglycerol-LD-TAG pool (DAG), inositol, acyl-CoA and acyl-glycerol. **Table S2** contains citations of experimental studies that have validated these antimetabolites. The structural analogs of these metabolites can be used for the treatment of HCV-associated liver disorders. Additionally, we identified sn-glycerol-3-phosphate, which is the precursor of the Kennedy pathway required for the *de novo* biosynthesis of glycolipids, such as phosphatidate (PA), DAG, TAG from dihydroxyacetone phosphate (DHAP) and acyl-CoA. Because of the similarity between the HCV lipid composition and VLDL/LDL⁴², we allowed the hepatocyte model to consume the VLDL, LDL and HDL remnants. This decision is consistent with that of Saito et al⁴³, who reported that HCV assembly was not affected by the addition of LDL to the growth medium of the Huh 7.5.1-8 cell line when the squalene synthase gene was downregulated. Our model-based simulations predicted the effect of consuming the remnants of all of the lipid metabolites, including the metabolites involved in cholesterol biosynthesis. However, we found that a structural analog of DAG prevented HCV assembly even in the presence of these lipid remnants. Herker et al⁴⁴ reported that the knockout of *DGATI* but not *DGAT2* reduced HCV assembly, and Hirata et al⁴⁵ reported that downregulation of *SGMS1* and *SGMS2* reduced HCV replication; all of the enzymes encoded by these genes use DAG as a substrate. When we constrained the model to consume the remnants of chylomicrons that included 90% of the TAGs and the remnants of the VLDL, LDL and HDL, DAG reduced the HCV assembly by 19%. Based on our model simulations, we observed that targeting DAG

1 reduced the HCV assembly even when the hepatocytes consumed the VLDL, LDL and HDL
2 remnants.

3 **Integrative analysis**

4 The integration of GEMs with transcriptomics data may reveal metabolic alterations in response
5 to HCV and eventually lead to the identification of candidate biomarkers for the progressive
6 stages (cirrhosis, dysplastic nodule, and early and advanced HCC). First, we downloaded the
7 microarray data series GSE6764 from the Gene Expression Omnibus public repository.⁴⁶ The data
8 set contains gene expression data from liver samples obtained from 10 healthy subjects, 18
9 cirrhotic patients, 14 dysplastic nodule patients, 17 early HCC patients and 18 advanced HCC
10 patients. These patients were infected only with HCV, and none of the patients had other hepatic
11 diseases. We used the PIANO R package⁴⁷ to perform a pairwise comparison of the gene
12 expression for the cirrhosis, dysplastic nodule, early HCC and advanced HCC samples with that
13 of the control samples. We also used PIANO to run the Reporter Features algorithm⁴⁸ using Gene
14 Ontology (GO) Biological Process (BP) terms to reveal the global biological changes in the
15 occurrence of each stage (**Supplementary Data 2**).

16

17 We also used the R platform to analyze the CNV, DNA methylation and microRNAs from the
18 TCGA samples (**Supplementary Data 3**). These additional omics data may provide further
19 explanations for the changes in gene expression identified by the Reporter Metabolite and
20 Reporter Subnetwork analyses. The CNV indicates the genomic regions that present a gain and
21 loss of DNA.⁴⁹ DNA methylation measures the methylation of CpG-dinucleotides on a genome
22 scale.⁵⁰ Methylation in the promoter region is negatively correlated with gene expression,
23 whereas methylation within the open reading frame is negatively or positively correlated with
24 gene expression.^{51, 52} Finally, the upregulation of microRNAs leads to the downregulation of the

1 target gene.⁵³ We restricted our analysis to samples classified as tumor degrees I, II, III, and VI
2 and normal samples (**Supplementary Data 4**). We used a MATLAB GUI to run the Reporter
3 Metabolite and Reporter Subnetwork algorithms (**See Method section**). Our GUI indicated the
4 metabolites that presented significant (P -value < 0.05) changes in metabolism for the four
5 different progressive stages (**Figure S1C, Supplementary Data 5**). We exported the
6 subnetwork for each stage to graphs saved as an XGMML file that could be opened in Cytoscape
7 (**Figure S1D, Supplementary Data 6**). Finally, we merged the CNV analysis with the
8 reported reactions from the Reporter Subnetwork analysis of the early HCC stage
9 (**Supplementary Data 7**).

10 **Keratan sulfate metabolism is altered in the cirrhosis stage**

11 Our results indicated that 17% of the identified Reporter Metabolites were involved in the
12 biosynthesis and degradation of sulfate proteoglycans. The metabolism associated with keratan
13 and chondroitin sulfate was significantly changed in the cirrhosis stage, whereas the metabolism
14 associated with keratan sulfate was significantly changed in the dysplastic nodule and early HCC
15 stages. The Reporter Features analysis depicts the metabolic changes in the hyaluronan catabolic
16 process (GO:0030214) in cirrhosis. The non-sulfate proteoglycan hyaluronic acid was
17 significantly increased in the sera of the HCV patients with cirrhosis compared with that of the
18 HCV patients without cirrhosis.⁵⁴ Additionally, lumican, which is a member of the family of
19 keratan sulfate proteoglycans, was elevated in the HCV patients.⁵⁵ In our gene expression data,
20 lumican (*LUM*) was upregulated (fold change = 1.23 and adjusted P -value < 0.02) during the
21 cirrhosis stage. Therefore, we propose that the chondroitin and keratan sulfate levels may be used
22 as candidate biomarkers for cirrhosis resulting from HCV infection.

Acyl-carnitine metabolism is altered in the dysplastic nodule stage

Based on the Reporter Metabolite analysis, we found that the metabolism associated with acyl-carnitine was significantly changed in the dysplastic nodule and early HCC stages. Additionally, we identified O-propanoylcarnitine as a Reporter Metabolite in early HCC and advanced HCC and O-butanoylcarnitine as a Reporter Metabolite only in advanced HCC (**Figure S2**). This finding is consistent with metabolomics studies of serum and tumor samples obtained from HCC patients. Thus, the oleoylcarnitine, palmitoylcarnitine, linoleic-carnitine and dodecanoylcarnitine levels are significantly changed in the sera of HCC patients infected with HCV.⁵⁶ O-Propanoylcarnitine (with betaine) has been proposed as a biomarker for the detection of early HCC.⁵⁷ Our analysis showed that long-chain acyl-carnitine in the endoplasmic reticulum was a Reporter Metabolite for the dysplastic nodule and early HCC stages (**Figure S2**). Therefore, these metabolites could be used as biomarkers for the dysplastic nodule stage based on their levels in the blood or urine. Our analysis of altered O-propanoylcarnitine levels was supported by measurements performed in the early HCC stage.⁵⁷

The Reporter Subnetwork analysis showed the role of *SLC22A5* in the translocation of acyl-carnitine into the ER. *SLC22A5* also transported carnitine from the extracellular space to the cytosol (**Figure 1B**) and was upregulated in the early and advanced HCC stages. Next, we searched for other genes associated with these metabolites. The *CRAT* gene was downregulated in the early and advanced HCC stages (**Figure 1B**). Furthermore, we identified 13 genes involved in the carnitine shuttle employed in the liver GEM (**Supplementary Data 8**). *SLC22A1*, *SLC25A20* and *CPT2* were all downregulated in the early or advanced HCC stages, whereas *CPT1A* was upregulated only in the advanced HCC stage.

1 ***BCAT1* gene and collagen effect on carnitine biosynthesis**

2 The Reporter Metabolite analysis indicated that carnitine metabolism was changed in the early
3 HCC stage and gamma-butyrobetaine and 3-hydroxy-N6,N6,N6-trimethyl-L-lysine metabolism
4 was significantly changed in the cirrhosis stage. This result is consistent with a metabolomics
5 study by Zhou et al⁵⁸, who reported that the carnitine level was significantly changed in the sera
6 of HCC and cirrhotic patients compared with healthy subjects and that the carnitine level was
7 significantly different between the HCC and cirrhotic patients. The authors reported that the
8 downregulation of carnitine might be related to the essential amino acids lysine and methionine,
9 which are carnitine precursors; however, our integrative analysis provided another possible
10 mechanistic explanation for this finding: the upregulation of *BCAT1* and *PLOD3* and
11 downregulation of *BBOX1* might have an effect on the downregulation of carnitine in early HCC
12 (see below).

13

14 The Reporter Subnetwork analysis demonstrated the downregulation of *BBOX1* (the last gene in
15 the carnitine biosynthesis pathway) in early and advanced HCC (**Figure 1B**). The DNA
16 methylation probes of *BBOX1* showed the hypomethylation of the non-CpG island with a delta
17 beta value = 0.1 in the promoter region (**Figure 1C and Figure S3**), which may have caused
18 *BBOX1* over-expression, although the gene was downregulated at the RNA level. Therefore, we
19 searched for a microRNA that targets *BBOX1* in the miRanda and TargetScan (release 6.0)
20 databases. We found hsa-miR-222 (hsa-miR-222-3p in release 7.0) in the target list based on the
21 network constructed by Gu et al,⁵⁹ who developed a systems biology approach to reconstruct
22 networks of genes, transcription factors and microRNAs.⁵⁹ This microRNA was reported to be
23 intensively over-expressed between normal subjects and HCC patients⁶⁰. We found that hsa-miR-

222 was also upregulated in the TCGA samples (**Figure 1E**). Therefore, the overall expression of *BBOX1* may be affected by the over-expression of hsa-miR-222.

The Reporter Subnetwork analysis also showed that alpha-ketoglutarate (AKG) was related to the largest number of downregulated reactions, including that of *BBOX1*; however, one reaction for *BCAT1* was upregulated in the early and advanced HCC stages. *BCAT1* converts branched amino acids (valine, leucine and isoleucine) and AKG to their corresponding keto-acids and glutamate and thus affects AKG metabolism. The DNA methylation analysis showed that *BCAT1* had a hypermethylated CpG island within the gene body (**Figure 1D and Figure S3**). Additionally, we integrated all of the reactions containing AKG with the observed gene expression (**Supplementary Data 9**) and found that only *BCAT1* and *PLOD3* were upregulated in HCC. *PLOD3* is the precursor of collagen biosynthesis and catalyzes the conversion of the carnitine biosynthesis precursor [protein]-L-lysine to procollagen 5-hydroxy-L-lysineuses (**Figure 1B**). Ura et al⁶¹ reported that the microRNA hsa-miR-187 targets *PLOD3* and found that other genes were also decreased in HCC relative to chronic hepatitis B and C. Ura et al⁶¹ evaluated the target microRNAs predicted using MIRANDA Pro3.0 in liver tissues obtained from 14 HCV and 12 HBV HCC patients against 9 normal liver tissues. Taken together, these results suggest that *PLOD3* upregulation may lead to the depletion of [protein]-L-lysine, which is required for carnitine biosynthesis, and the downregulation of *BBOX1* may lead to the depletion of synthesized carnitine in early and advanced HCC.

1 SAME metabolism is affected by the amplification of gain regions on chromosome
2 1q

3 We identified S-adenosylmethionine (SAME) and S-adenosylhomocysteine (SAH) as the top
4 Reporter Metabolites. We also identified SAME metabolism using the Reporter Subnetwork
5 algorithm (**Figure 2A**). SAME is biosynthesized from methionine and ATP by the methionine
6 adenosyl transferase genes (*MAT1A*, *MAT2A* and *MAT2B*). *MAT1A* is significantly
7 downregulated in early HCC and advanced HCC. The role of SAME in hepatic diseases and HCC
8 was reviewed by Lu and Mato.⁶² One explanation for the downregulation of *MAT1A* could be the
9 upregulation of the microRNAs hsa-miR-664, hsa-miR-495, and hsa-miR-485-3p.⁶³ Yang et al
10 found these microRNAs by searching for microRNAs targeting *MAT1A* in miRanda and
11 TargetScan and validated the key roles of these microRNAs in liver cancer cell lines.⁶³ In the
12 TCGA database, hsa-miR-664 is upregulated in liver cancer samples (**Figure 2B**). In the second
13 reaction of SAME metabolism, SAME is converted to SAH by a methyltransferase. The Reporter
14 Subnetwork analysis reported three downregulated genes in the early and advanced HCC stages
15 related to methyltransferase: *GNMT*, which converts glycine to sarcosine; *SARDH*, which
16 catalyzes the conversion of dimethylglycine to sarcosine; and *PENT*, which catalyzes the
17 conversion of phosphatidylethanolamine to phosphatidylcholine.

18
19 The integration of transcriptomics data with the GEM was used to identify other upregulated
20 methyltransferase genes that could use SAME. To identify the highly expressed methyltransferase
21 genes in the model, we collected the reactions containing SAM and mapped these reactions to
22 their gene expression profiles and the CNV analysis (**Supplementary Data 10**). This analysis
23 identified the genes *ASH1L*, *METTL13*, *SMYD2*, *TARBP1* and *SMYD3* as upregulated and
24 amplified. In addition, hsa-miR-664 was amplified and located on the same chromosome (1q) as

these genes (**Figure 2A and Figure 2B**). This analysis identified six genes that were upregulated (*SMYD3*, *ASHIL*, *SETDB2*, *SUV420H1*, *WHSC1* and *MLL5*) and five genes that were downregulated (*DOTIL*, *MLL4*, *MLL*, *WHSC1L1* and *MLL3*), and these downregulated genes were involved in SAME and carnitine biosynthesis. Kusano et al⁶⁴ reported amplification in chromosome 1q in 78% of HCC cases, and Skawran et al⁶⁵ reported the upregulation of genes on chromosome 1q using microarray analysis. Therefore, this amplification could affect SAME metabolism. The remaining genes (*ASHIL*, *METTL13*, *SMYD2*, *TARBP1* and *SMYD3*) encode enzymes that may be involved in alternative uses of SAME.

Through the integration of the CNV analysis with the Reporter Subnetwork analysis, we identified two genes (*GNPAT* and *PPOX*) that were amplified and located on chromosome 1q (**Figure 2A and Figure S4**). *PPOX* may lead to an increase in ROS through the heme pathway because *NOX4* is upregulated in advanced HCC (**Figure S5A and Supplementary Data 11**), and *GNPAT* could affect acyl-CoA metabolism as discussed below.

Finally, the Reporter Subnetwork analysis demonstrated that methionine was linked to *SLC7A6*. This gene transports amino acids from the cytoplasm to the extracellular space (and vice versa) and is associated with major amino acid transport reactions in the model (**Supplementary Data 12**). The Reporter Subnetwork analysis showed that *SLC7A6* increases the cytoplasmic citrate level by transporting additional citrate from the extracellular space to the cytoplasm (**Figure S5B**). This result is consistent with metabolomics studies that observed significantly decreased citrate levels in the sera of HCC patients.¹⁹

1 **Role of *GNPAT* in acyl-CoA metabolism**

2 We identified DHAP and sn-glycerol-3-phosphate as Reporter Metabolites in early and advanced
3 HCC, and metabolomics studies have shown that n-glycerol-3-phosphate is downregulated in
4 these stages.¹⁹ The *GPD1* gene, which catalyzes DHAP to glycerol-3-phosphate, was
5 downregulated in early and advanced HCC, whereas the *GPD1L* and *GPD2* genes were
6 unregulated in early HCC but upregulated in advanced HCC according to the Reporter
7 Subnetwork analysis (**Figure 3**). Furthermore, *GNPAT* that encodes a peroxisomal enzyme that
8 catalyzes the conversion of both DHAP and long-chain acyl-CoA to 1-acyl-DHAP, was
9 upregulated in early and advanced HCC. The CNV analysis showed that *GNPAT* was amplified
10 and located on chromosome 1q. Therefore, the upregulation of *GNPAT* and downregulation of
11 *GPD1* may decrease the amount of DHAP converted to sn-glycerol-3-phosphate in the early and
12 advanced HCC stages.

13 Next, we examined the other reactions involved in lipid biosynthesis in the peroxisome and found
14 that the reaction transporting long-chain acyl-CoA from the cytosol to the peroxisome was
15 upregulated in early and advanced HCC (**Supplementary Data 11**). The *BCAP31* gene was
16 upregulated in early and advanced HCC, whereas the other genes associated with this reaction
17 were upregulated only in advanced HCC. The *AGPS* gene, which catalyzes the conversion of both
18 fatty alcohol and 1-acyl-DHAP to 1-alkyl-DHAP, was upregulated only in advanced HCC,
19 whereas the *FAR1* and *FAR2* genes were unregulated in early and advanced HCC. In the third
20 reaction, DHAP reductase converts 1-alkyl-DHAP to lysophosphatidic acid (LPA); however, the
21 associated gene for this enzyme has not been identified because it is not in sequence databases,
22 such as UniProt. Our analysis could explain the endogenous biosynthesis of LPA from DHAP in
23 early and advanced HCC, and the composition of ether lipids in LPA could be altered in early
24 HCC. Furthermore, our analysis indicated another pathway that consumes acyl-CoA. The

upregulation of *BCAP31* leads to the increased transport of acyl-CoA to the peroxisomes, and *GNPAT* can use more acyl-CoA for the synthesis of ether lipids. This pathway may be interesting because the knock-out of *AGPAT1* in mice delayed the occurrence of HCC due to the increased availability of acyl-CoA.⁶⁶ Furthermore, of the genes that use acyl-CoA, only the *AGPAT1* gene is upregulated (**Figure 3**).

Discussion

In addition to playing a role in HCV drug development and new biomarker identification⁹, systems biology also contributes to viral research.⁶⁷ Here, we demonstrated how the FBA method combined with the antimetabolite concept could be applied to identify host target metabolites that impede the HCV assembly reaction. We identified metabolites involved in sterol and sphingolipid biosynthesis and phosphoinositides, and our results are consistent with that of experimental studies (**Table S2**). When we constrained the model to consume LDL, VLDL, and HDL remnants, we found that the structural analogs of DAG could be used as antimetabolites. We predicted that targeting the metabolite DAG pool could avoid the effect of consuming these remnants, whereas targeting all of the other lipid metabolites affects the uptake of these remnants.

We also integrated transcriptomics data for the progressive stages caused by HCV infection (cirrhosis, dysplastic nodule, and early and advanced HCC) with the GEM *iHepatocytes2322*. We identified acyl-carnitine metabolism as significantly changed in the dysplastic nodule and early HCC stages. Altered acyl-carnitine levels have also been reported in metabolomics studies.¹⁹ Our analysis predicted that acyl-carnitine metabolism was also changed in the dysplastic nodule stage and provided an explanation for the downregulation of carnitine in early HCC. The upregulation

of *PLOD3* can result in the use of [protein]-L-lysine (the precursor of carnitine) for collagen biosynthesis, and the downregulation of *BBOX1* can reduce the endogenous biosynthesis of carnitine. Finally, we found that the *BCAT1* and *PLOD3* genes involved in AKG metabolism were upregulated only in the early and advanced HCC stages, and this could further reduce the flux of AKG towards carnitine (**Figure 1B**).

Chromosome instability and rearrangement may be factors underlying HCC,⁶⁸ and regions of chromosome 1q have been reported to be amplified in HCC.⁶⁵ We analyzed the CNV from TCGA between cancer and normal samples and found that SAME metabolism was affected by genome rearrangements on chromosome 1q (**Figure 2A**). First, hsa-miR-664 is located on chromosome 1q and targets the *MAT1A* gene⁶³ (**Figure 2B**), which may decrease SAME. Additionally, the expression of *ASH1L*, *METTL13*, *SMYD2*, *TARBP1* and *SMYD3*, which encode enzymes that consume SAME, was amplified. Our analysis suggests that the *GNPAT* gene is both amplified and upregulated, which may explain the downregulation of sn-glycerol-3-phosphate and long acyl-carnitine observed in metabolomics studies.¹⁹ Taken together, the decrease in long-chain acyl-carnitine could be related to the upregulation of *AGPAT1*, *GNPAT*, *BCAP31* and *SLC25A5*. The first three genes utilize a greater amount of acyl-CoA (**Figure 3**), whereas *SLC22A5* transports acyl-carnitine into the ER (**Figure 1B**).

In conclusion, our integrative analysis of HCV-induced HCC identified new candidate drug targets and several new candidate biomarkers that required further validation through dedicated

1 clinical studies. Importantly, our analysis provides a mechanistic explanation for changes
2 observed in lipid and carnitine metabolism that are associated with early HCC.

3 **Methods**

4 **Flux balance analysis**

5 The GEM includes a list of chemical reactions that can be represented as a stoichiometric matrix
6 S .^{69, 70} Each row represents a metabolite, and each column represents a flux value for a reaction.

7 Each element S_{ij} , represents the coefficient of metabolite i th in the reaction j th, and the vector V
8 is a set of all fluxes. FBA maximizes a reaction that refers to a growth or biomass in the steady
9 state condition ($S \cdot V = 0$).^{69, 70} In our case, FBA maximizes the HCV assembly reaction v_{HCV}
10 (Equation 1) where UB and LB are the upper and lower bound values for each flux, respectively.

11 In the antimetabolite technique, we first determine the set of reactions that has a metabolite as a
12 substrate or product, and we then set the lower and upper bounds for these fluxes equal to zero.

13 Finally we solve the FBA problem in Equation 1 and report the change in v_{HCV} , using the
14 RAVEN toolbox⁷¹ with the MOSEK solver (Free Academic License). The metabolite is essential
15 for the HCV assembly if $v_{HCV} = 0$.

$$\begin{array}{ll}
 \text{Max} & v_{HCV} \\
 \text{S.t.} & \\
 16 & S \cdot V = 0 \\
 & V \leq UB \\
 & -V \leq -LB
 \end{array} \quad (1)$$

1 **HCV transcriptomics data**

2 We used the liver tissue microarray series GSE6764 from the Gene Expression Omnibus public
3 repository (GEO). The data consist of 75 samples from 48 patients at different states as follows:
4 normal liver (n = 10), cirrhosis (n = 13), dysplastic nodules (n = 17), early HCC (n = 18) and
5 advanced HCC (n = 17). The samples were obtained from patients undergoing resection or liver
6 transplants. The patients had no other liver diseases, including HBV-positive markers, alcohol
7 consumption history, steatohepatitis, hemochromatosis, and other causes of chronic liver disease.
8 The samples were classified by expert pathologists according to the criteria of the International
9 Working Party.⁷² The 10 control samples were collected from patients undergoing resection for
10 hepatic hemangioma (1), focal nodular hyperplasia (3), adenoma/cystadenoma (2),
11 neuroendocrine tumor (1), and living donor liver transplantation (1).¹⁶
12 We used the R package PIANO⁴⁷ to analyze and normalize the data and identify differentially
13 expressed genes. We used the false discovery rate (FDR) method to calculate the adjusted *P*-
14 values and avoid false significance tests resulting from a large number of performed tests.⁷³ We
15 analyzed the differential gene expression in the four different stages (cirrhosis, dysplasia, early
16 and advance HCC) compared with that of the normal samples.

17

18 **Updating the model with transcriptomics data**

19 To include genes that are absent from the model as well as additional genes from the
20 transcriptomics analysis, we created a graphical user interface (GUI) in MATLAB based on the
21 RAVEN toolbox⁷¹ to perform this task and facilitate running both the Reporter Metabolite and
22 Reporter Subnetwork algorithms (**Supplementary Data 13**). The Reporter Metabolite
23 algorithm reconstructs the associations of a metabolite gene and calculates a *Z* score for each
24 metabolite using the adjusted *P*-values of the genes connected to this metabolite (**Figure S1C**).

1 We considered the metabolism of a metabolite to be significantly changed if the P -value was <
2 0.05. The GUI inputs were as follows: *iHepatocytes2322* as the core model, HMR 2.0 as the
3 reference model, proteins from a proteomics study of the livers of 15 HCV-infected patients⁷⁴
4 (**Supplementary Data 14**), and four files for each independent pairwise comparison. Our GUI
5 output all of the Reporter Metabolite data into one Excel file, where each stage was stored in a
6 sheet. The Reporter Subnetwork algorithm used gene expression data to extract the highly
7 affected subnetworks as a set of metabolic reactions affected by the differentially expressed genes
8 between two stages (**Figure S1D**). We used our GUI to remove the highly connected metabolites
9 (i.e., cofactors, H₂O and other metabolites listed in **Table S3**) from the model. Then, our GUI
10 uploaded the modified model and the adjusted P -value for each gene in the model for each stage
11 to the Biomet database, and then the Reporter Subnetwork algorithm was executed ten times. The
12 GUI selected the best subnetwork with the maximum Z score and exported it in XGMML format
13 so that it could be opened easily with Cytoscape. Finally, we applied the Reporter Features
14 algorithm,⁴⁸ which is a gene-set enrichment algorithm that replaces the metabolites uses genes
15 associated with GO biological process terms and calculates the Z score for a GO term from the
16 adjusted P -value of the probes associated with this GO term. We used the R package biomaRt to
17 annotate the probes with their corresponding GO terms and the PIANO function *runGSA* to
18 execute the Reporter Features algorithm. To isolate the metabolic GO terms, we used the R
19 library GO.db to identify all of the biological process children of the "GO:0008152 metabolic
20 process," which resulted in 9229 metabolic GO terms. We only reported the changed metabolic
21 GO terms that had "GO:0008152 metabolic process" as the parent and a non-directional P -value
22 < 0.05.

23

1 **Copy number variation, DNA methylation and microRNA from TCGA**

2 We downloaded the analyzed data with annotations for the liver cancer samples (level 3) for
3 CNVs, DNA methylation and microRNA from the TCGA web site. For CNVs, we used the
4 Biomart service to identify the start and end points for Ensembl genes in the HMR 2.0 model. For
5 each sample mapped to the genome (h19), we assigned the segment mean value to a HMR 2.0
6 gene that overlapped with a segment. Finally, we selected the genes that had absolute segment
7 values greater than 0.15 in more than 50% of the cancer samples and average *segment_mean*
8 values of the cancer samples that differed from the average *segment_mean* values of the normal
9 samples in the non-parametric test (wilcox.test) with a P -value $< 0.01^{75}$.

10 For DNA methylation, we built two arrays containing the beta value for the probe and the average
11 beta values for the probes in the island. We calculated the delta beta value as the absolute value
12 for the difference in the average beta values between the cancer and normal samples. We used
13 *HumanMethylation450 v1.2* from the Illumina web site to annotate the CpG island and gene
14 names. We used the PIANO package to calculate the adjusted P -values from the t-test. We
15 selected the probes or the CpG islands with delta beta values ≥ 0.1 and adjusted P -values < 0.05 ,
16 and we classified a probe or CpG island as hypomethylated if the average beta value for the
17 cancer sample was less than the average beta value for the normal sample or hypermethylated if
18 the average beta value for the cancer sample was greater than the average beta value for the
19 normal sample.

20 Finally, for the microRNA analysis, we built an array from the reads per million for each
21 microRNA⁷⁶ and used the default commands in the edgeR R package to identify the expressed
22 microRNAs.⁷⁷

1 Acknowledgements

2 We thank Francesco Gatto for his help in analyzing the TCGA data. Additionally, we thank Leif
3 Väre for his help in analyzing the gene expression data and Reporter Features algorithm.
4 Ibrahim Elsemman acknowledges the Egyptian Cultural Affairs and Missions Sector for their
5 financial support. We also acknowledge funding from the Knut and Alice Wallenberg
6 Foundation.

7 Author contributions

8 IE analyzed all of the data, developed the methods, incorporated *iHCV* into the *iHepatocytes2322*
9 GEM, performed the simulations and drafted the manuscript. AM assisted with the clinical data.
10 SS assisted with the reconstruction of the *iHCV*. AM, THS and JN supervised the work and edited
11 the manuscript. JN conceived and designed the project. All of the authors read and approved the
12 final manuscript.

13 Conflicts of interest

14 The authors declare that there are no conflicts of interest related to this paper.

15 Figure Legends

16

17 **Figure 1. Effective approach describing the metabolic alterations for each stage caused by**
18 **HCV infection. (A)** HCV has three structural proteins: the core and two envelope proteins *E1*
19 and *E2*. HCV contains lipids and the *apoE* and *apoB* lipoproteins. The HCV structure is
20 converted to a biosynthesis reaction, where the coefficients a_1 , a_2 , a_3 , a_4 and a_5 represent the ratio
21 of the *E1*, *E2* and core proteins; lipids; and HCV particles to the amount of HCV RNA. This

1 reaction is integrated with the hepatic model *iHepatocytes2322*. The new model can be used to
2 identify metabolites crucial for HCV assembly via antimetabolites and can be integrated with
3 transcriptomics studies (GSE6764) using the GEO database to identify significantly changed
4 metabolites and reactions. Moreover, the genes reported in the Reporter Subnetwork analysis
5 were mapped with copy number variations of liver cancer from TCGA. **(B)** Carnitine metabolism
6 is integrated with gene expression in early HCC. The blue, red and black arrows indicate that the
7 gene(s) associated with this reaction are downregulated, upregulated and unregulated (adjusted *P*-
8 value > 0.05), respectively. The reaction is also unregulated if it involves both upregulated and
9 downregulated genes. **(C)** Boxplots show the hypomethylated DNA methylation levels for non-
10 promoter CpG islands (NPI) of *BBOX1* in the HCC and normal TCGA samples, where the delta
11 beta value = 0.1 (adjusted *P*-value < 0.05). **(D)** Boxplots show the hypermethylated DNA
12 methylation levels for body CpG islands (BI) of *BCAT1* in both the HCC and normal TCGA
13 samples, where the delta beta value = 0.1 (adjusted *P*-value < 0.05). **(E)** Boxplots show the over-
14 expression level of hsa-miR-222 in the HCC and normal TCGA samples (adjusted *P*-value <
15 0.05).

16

17 **Figure 2. SAME metabolism is integrated with copy number variations on chromosome 1q**
18 **from liver cancer in TCGA.** **(A)** Blue, red and black arrows indicate that the gene(s) associated
19 with this reaction are downregulated, upregulated and unregulated (adjusted *P*-value > 0.05),
20 respectively. The reaction is also unregulated if it involves both upregulated and downregulated
21 genes. A small red rectangle in chromosome 1q indicates that this gene or microRNA is
22 amplified. **(B)** Boxplots show the over-expression level of hsa-miR-664 in the HCC and normal
23 TCGA samples (adjusted *P*-value < 0.05).

Figure 3. Acyl-CoA metabolism is integrated with gene expression in early HCC. Blue, red and black arrows indicate that the gene(s) associated with this reaction are downregulated, upregulated and unregulated (adjusted P -value > 0.05), respectively. The reaction is also unregulated if it involves both upregulated and downregulated genes. The dashed orange line indicates that the DHAP reductase gene has not been annotated for human lines. *ACS*, acyl-CoA synthase; LPA, lysophosphatidic acid; PA, phosphatidate; DAG, diacylglycerolipids; and TAG, triglycerolipid.

References

1. T. K. H. Scheel and C. M. Rice, *Nat Med*, 2013, **19**, 837-849.
2. A. L. Mason, J. Y. Lau, N. Hoang, K. Qian, G. J. Alexander, L. Xu, L. Guo, S. Jacob, F. G. Regenstein, R. Zimmerman, J. E. Everhart, C. Wasserfall, N. K. Maclaren and R. P. Perrillo, *Hepatology*, 1999, **29**, 328-333.
3. A. A. Butt, W. Xiaoqiang, M. Budoff, D. Leaf, L. H. Kuller and A. C. Justice, *Clinical Infectious Diseases*, 2009, **49**, 225-232.
4. Q. Choo, G. Kuo, A. Weiner, L. Overby, D. Bradley and M. Houghton, *Science*, 1989, **244**, 359-362.
5. T. J. Liang, *Nat Med*, 2013, **19**, 869-878.
6. D. Ge, J. Fellay, A. J. Thompson, J. S. Simon, K. V. Shianna, T. J. Urban, E. L. Heinzen, P. Qiu, A. H. Bertelsen, A. J. Muir, M. Sulkowski, J. G. McHutchison and D. B. Goldstein, *Nature*, 2009, **461**, 399-401.
7. J.-M. Pawlotsky, *Hepatology*, 2011, **53**, 1742-1751.
8. R. Singaravelu, N. Nasheri, A. Sherratt and J. P. Pezacki, *Hepatology*, 2012, **56**, 1-4.
9. R. C. Ireton and M. Gale, Jr., *Current topics in microbiology and immunology*, 2013, **363**, 143-167.
10. A. L. Rasmussen, D. L. Diamond, J. E. McDermott, X. Gao, T. O. Metz, M. M. Matzke, V. S. Carter, S. E. Belisle, M. J. Korth, K. M. Waters, R. D. Smith and M. G. Katze, *Journal of virology*, 2011, **85**, 11646-11654.
11. A. Forner, J. M. Llovet and J. Bruix, *The Lancet*, **379**, 1245-1255.
12. A. Forner and J. Bruix, *The Lancet Oncology*, 2012, **13**, 750-751.
13. J. Bruix and M. Sherman, *Hepatology*, 2005, **42**, 1208-1236.
14. D. C. Rockey, S. H. Caldwell, Z. D. Goodman, R. C. Nelson and A. D. Smith, *Hepatology*, 2009, **49**, 1017-1044.
15. G. Nathalie, D. Delphine, D. Françoise, J. Jean Louis, V. Dominique, B. Jacques, B. Pierre and P. Valerie, *Laboratory Investigation*, 2006, **86**, 951-958.

1 16. E. Wurmbach, Y.-b. Chen, G. Khitrov, W. Zhang, S. Roayaie, M. Schwartz, I. Fiel, S. Thung,
2 V. Mazzaferro, J. Bruix, E. Bottinger, S. Friedman, S. Waxman and J. M. Llovet,
3 *Hepatology*, 2007, **45**, 938-947.
4 17. R. N. Aravalli, C. J. Steer and E. N. Cressman, *Hepatology*, 2008, **48**, 2047-2063.
5 18. X. Wang, A. Zhang and H. Sun, *Hepatology*, 2013, **57**, 2072-2077.
6 19. D. Beyoglu and J. R. Idle, *J Hepatol*, 2013, **59**, 842-858.
7 20. A. Mardinoglu and J. Nielsen, *J Intern Med*, 2012, **271**, 142-154.
8 21. A. Mardinoglu and J. Nielsen, *New paradigms for metabolic modeling of human cells*,
9 *Curr Opin Biotechnol*. 2015 Jan 2;34C:91-97. doi: 10.1016/j.copbio.2014.12.013.
10 22. S. Shoaie and J. Nielsen, *Front Genet*, 2014, **5**.
11 23. I. Thiele and B. O. Palsson, *Nat. Protocols*, 2010, **5**, 93-121.
12 24. M. Kanehisa, S. Goto, S. Kawashima, Y. Okuno and M. Hattori, *Nucleic Acids Research*,
13 2004, **32**, D277-D280.
14 25. A. Mardinoglu, F. Gatto and J. Nielsen, *Biotechnology journal*, 2013, **8**, 985-996.
15 26. K. R. Patil and J. Nielsen, *Proceedings of the National Academy of Sciences of the United*
16 *States of America*, 2005, **102**, 2685-2689.
17 27. H. U. Kim, S. Y. Kim, H. Jeong, T. Y. Kim, J. J. Kim, H. E. Choy, K. Y. Yi, J. H. Rhee and S. Y.
18 Lee, *Molecular Systems Biology*, 2011, **7**, n/a-n/a.
19 28. H. U. Kim, T. Y. Kim and S. Y. Lee, *Molecular BioSystems*, 2010, **6**, 339-348.
20 29. P. Ghaffari, A. Mardinoglu, A. Asplund, S. Shoaie, C. Kampf, M. Uhlen and J. Nielsen, *Sci.*
21 *Rep.*, 2015, **5**.
22 30. P.-J. Kim, D.-Y. Lee, T. Y. Kim, K. H. Lee, H. Jeong, S. Y. Lee and S. Park, *Proceedings of the*
23 *National Academy of Sciences*, 2007, **104**, 13638-13642.
24 31. A. Mardinoglu, R. Agren, C. Kampf, A. Asplund, M. Uhlen and J. Nielsen, *Nat Commun*,
25 2014, **5**.
26 32. I. Thiele and B. Palsson, *Nat Protoc*, 2010, **5**, 93 - 121.
27 33. R. Jain and R. Srivastava, *BMC Systems Biology*, 2009, **3**, 121.
28 34. E. W. Birch, N. A. Ruggero and M. W. Covert, *PLoS Comput Biol*, 2012, **8**, e1002746.
29 35. F. Penin, J. Dubuisson, F. A. Rey, D. Moradpour and J.-M. Pawlotsky, *Hepatology*, 2004,
30 **39**, 5-19.
31 36. D. Quinkert, R. Bartenschlager and V. Lohmann, *Journal of virology*, 2005, **79**, 13594-
32 13605.
33 37. C. Scholtes, C. Ramière, D. Rainteau, L. Perrin-Cocon, C. Wolf, L. Humbert, M. Carreras,
34 A. Guironnet-Paquet, F. Zoulim, R. Bartenschlager, V. Lotteau, P. André and O. Diaz,
35 *Hepatology*, 2012, **56**, 39-48.
36 38. C. Chan, F. Berthiaume, K. Lee and M. L. Yarmush, *Biotechnology and Bioengineering*,
37 2003, **81**, 33-49.
38 39. L. Jerby, T. Shlomi and E. Ruppin, *Mol Syst Biol*, 2010, **6**, 56.
39 40. H. G. Wilcox and M. Heimberg, *Journal of Lipid Research*, 1987, **28**, 351-360.
40 41. R. Agren, A. Mardinoglu, A. Asplund, C. Kampf, M. Uhlen and J. Nielsen, *Mol Syst Biol*,
41 2014, **10**, 721.
42 42. A. Merz, G. Long, M.-S. Hiet, B. Brügger, P. Chlanda, P. Andre, F. Wieland, J. Krijnse-
43 Locker and R. Bartenschlager, *Journal of Biological Chemistry*, 2011, **286**, 3018-3032.
44 43. K. Saito, Y. Shirasago, T. Suzuki, H. Aizaki, K. Hanada, T. Wakita, M. Nishijima and M.
45 Fukasawa, *Journal of virology*, 2015, **89**, 2220-2232.

- 1 44. E. Herker, C. Harris, C. Hernandez, A. Carpentier, K. Kaehlcke, A. R. Rosenberg, R. V.
2 Farese and M. Ott, *Nat Med*, 2010, **16**, 1295-1298.
- 3 45. Y. Hirata, K. Ikeda, M. Sudoh, Y. Tokunaga, A. Suzuki, L. Weng, M. Ohta, Y. Tobita, K.
4 Okano, K. Ozeki, K. Kawasaki, T. Tsukuda, A. Katsume, Y. Aoki, T. Umehara, S. Sekiguchi,
5 T. Toyoda, K. Shimotohno, T. Soga, M. Nishijima, R. Taguchi and M. Kohara, *PLoS*
6 *Pathogens*, 2012, **8**, e1002860.
- 7 46. T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A.
8 Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L.
9 Robertson, N. Serova, S. Davis and A. Soboleva, *Nucleic Acids Research*, 2013, **41**, D991-
10 D995.
- 11 47. L. Väre, J. Nielsen and I. Nookaew, *Nucleic Acids Research*, 2013, **41**, 4378-4391.
- 12 48. A. Oliveira, K. Patil and J. Nielsen, *BMC Systems Biology*, 2008, **2**, 17.
- 13 49. C. Alkan, B. P. Coe and E. E. Eichler, *Nat Rev Genet*, 2011, **12**, 363-376.
- 14 50. P. A. Jones, *Nat Rev Genet*, 2012, **13**, 484-492.
- 15 51. X. Yang, H. Han, Daniel D. De Carvalho, Fides D. Lay, Peter A. Jones and G. Liang, *Cancer*
16 *Cell*, 2014, **26**, 577-590.
- 17 52. S. Kasinathan and S. Henikoff, *Cancer Cell*, 2014, **26**, 449-451.
- 18 53. P. T. Y. Law and N. Wong, *Journal of Gastroenterology and Hepatology*, 2011, **26**, 437-
19 449.
- 20 54. J. G. McHutchison, L. M. Blatt, M. De Medina, J. R. Craig, A. Conrad, E. R. Schiff, M. J.
21 Tong and T. C. I. S. Group, *Journal of Gastroenterology and Hepatology*, 2000, **15**, 945-
22 951.
- 23 55. A. Krishnan, X. Li, W.-Y. Kao, K. Viker, K. Butters, H. Masuoka, B. Knudsen, G. Gores and
24 M. Charlton, *Lab Invest*, 2012, **92**, 1712-1725.
- 25 56. J. F. Xiao, R. S. Varghese, B. Zhou, M. R. Nezami Ranjbar, Y. Zhao, T.-H. Tsai, C. Di Poto, J.
26 Wang, D. Goerlitz, Y. Luo, A. K. Cheema, N. Sarhan, H. Soliman, M. G. Tadesse, D. H.
27 Ziada and H. W. Ransom, *Journal of Proteome Research*, 2012, **11**, 5914-5923.
- 28 57. Q. Huang, Y.-x. Tan, P. Yin, G. Ye, P. Gao, X. Lu, H.-Y. Wang and G. Xu, *Cancer Research*,
29 2013, DOI: 10.1158/0008-5472.can-13-0308.
- 30 58. N. Qin, F. Yang, A. Li, E. Prifti, Y. Chen, L. Shao, J. Guo, E. Le Chatelier, J. Yao, L. Wu, J.
31 Zhou, S. Ni, L. Liu, N. Pons, J. M. Batto, S. P. Kennedy, P. Leonard, C. Yuan, W. Ding, Y.
32 Chen, X. Hu, B. Zheng, G. Qian, W. Xu, S. D. Ehrlich, S. Zheng and L. Li, *Nature*, 2014, **513**,
33 59-64.
- 34 59. Z. Gu, C. Zhang and J. Wang, *BMC Syst Biol*, 2012, **6**, 1752-0509.
- 35 60. C. Braconi, J. C. Henry, T. Kogure, T. Schmittgen and T. Patel, *Seminars in oncology*, 2011,
36 **38**, 752-763.
- 37 61. S. Ura, M. Honda, T. Yamashita, T. Ueda, H. Takatori, R. Nishino, H. Sunakozaka, Y. Sakai,
38 K. Horimoto and S. Kaneko, *Hepatology*, 2009, **49**, 1098-1112.
- 39 62. S. C. Lu and J. M. Mato, *S-adenosylmethionine in Liver Health, Injury, and Cancer*, 2012.
- 40 63. H. Yang, M. E. Cho, T. W. H. Li, H. Peng, K. S. Ko, J. M. Mato and S. C. Lu, *The Journal of*
41 *Clinical Investigation*, 2013, **123**, 285-298.
- 42 64. N. Kusano, K. Shiraishi, K. Kubo, A. Oga, K. Okita and K. Sasaki, *Hepatology*, 1999, **29**,
43 1858-1862.
- 44 65. B. Skawran, D. Steinemann, A. Weigmann, P. Flemming, T. Becker, J. Flik, H. Kreipe, B.
45 Schlegelberger and L. Wilkens, *Mod Pathol*, 2008, **21**, 505-516.

166. J. M. Ellis, D. S. Paul, M. A. Depetrillo, B. P. Singh, D. E. Malarkey and R. A. Coleman, *Toxicologic pathology*, 2012, **40**, 513-521.

167. G. L. Law, M. J. Korth, A. G. Benecke and M. G. Katze, *Nat Rev Micro*, 2013, **11**, 455-466.

168. M. Liu, L. Jiang and X.-Y. Guan, *Protein Cell*, 2014, **5**, 673-691.

169. J. M. Park, T. Y. Kim and S. Y. Lee, *Biotechnology Advances*, 2009, **27**, 979-988.

170. N. D. Price, J. A. Papin, C. H. Schilling and B. O. Palsson, *Trends in Biotechnology*, 2003, **21**, 162-169.

171. R. Agren, L. Liu, S. Shoaie, W. Vongsangnak, I. Nookaew and J. Nielsen, *PLoS Comput Biol*, 2013, **9**, e1002980.

172. *Hepatology*, 1995, **22**, 983-993.

173. W. S. Noble, *Nat Biotech*, 2009, **27**, 1135-1137.

174. D. L. Diamond, J. M. Jacobs, B. Paeper, S. C. Proll, M. A. Gritsenko, R. L. Carithers, Jr., A. M. Larson, M. M. Yeh, D. G. Camp, 2nd, R. D. Smith and M. G. Katze, *Hepatology*, 2007, **46**, 649-657.

175. F. Gatto, I. Nookaew and J. Nielsen, *Proceedings of the National Academy of Sciences*, 2014, **111**, E866-E875.

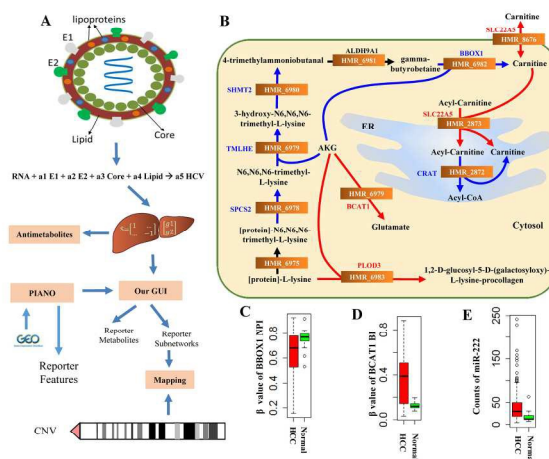
176. M. P. Hamilton, K. Rajapakshe, S. M. Hartig, B. Reva, M. D. McLellan, C. Kandoth, L. Ding, T. I. Zack, P. H. Gunaratne, D. A. Wheeler, C. Coarfa and S. E. McGuire, *Nat Commun*, 2013, **4**.

177. M. D. Robinson, D. J. McCarthy and G. K. Smyth, *Bioinformatics*, 2010, **26**, 139-140.

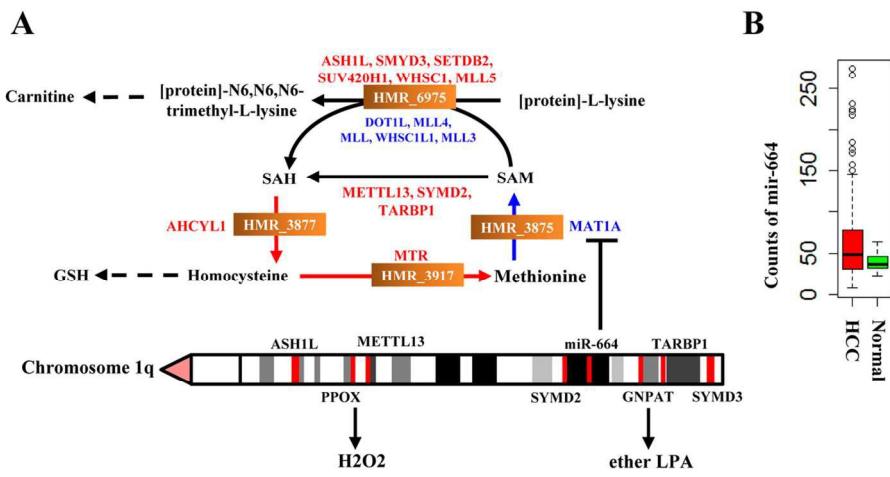
178. J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander and N. Schultz, *Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal*, 2013.

179. B. Phipson and A. Oshlack, *Genome Biology*, 2014, **15**, 465.

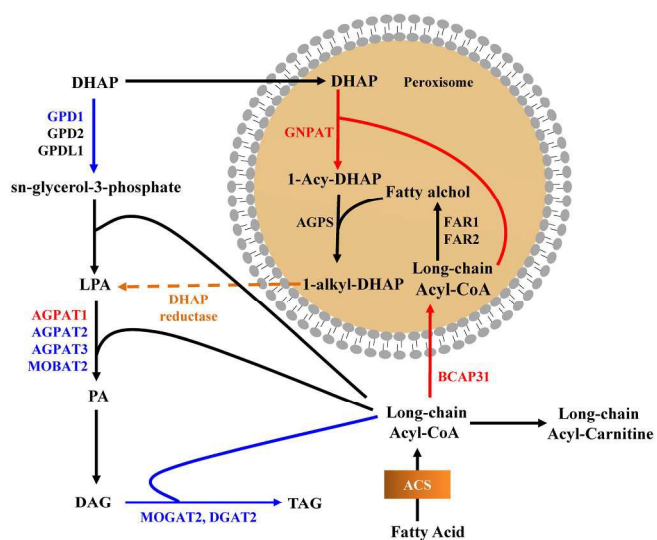
180. D. L. Diamond, J. M. Jacobs, B. Paeper, S. C. Proll, M. A. Gritsenko, R. L. Carithers, A. M. Larson, M. M. Yeh, D. G. Camp, R. D. Smith and M. G. Katze, *Hepatology*, 2007, **46**, 649-657.



209x148mm (300 x 300 DPI)



120x85mm (300 x 300 DPI)



209x148mm (300 x 300 DPI)